THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 1997 by The American Society for Biochemistry and Molecular Biology, Inc.

Vol. 272, No. 33, Issue of August 15, pp. 20456–20462, 1997
Printed in U.S.A.

Optimal Subsite Occupancy and Design of a Selective Inhibitor of Urokinase*

(Received for publication, May 12, 1997)

Song-Hua Ke‡, Gary S. Coombs§¶, Kathy Tachias‡, David R. Corey§‖, and Edwin L. Madison‡

From the ‡Department of Vascular Biology, The Scripps Research Institute, La Jolla, California 92037 and the §Department of Pharmacology and Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, Texas 75235

Human urokinase type plasminogen activator (u-PA) is a member of the chymotrypsin family of serine proteases that can play important roles in both health and disease. We have used substrate phage display techniques to characterize the specificity of this enzyme in detail and to identify peptides that are cleaved 840–5300 times more efficiently by u-PA than peptides containing the physiological target sequence of the enzyme. In addition, unlike peptides containing the physiological target sequence, the peptide substrates selected in this study were cleaved as much as 120 times more efficiently by u-PA than by tissue type plasminogen activator (t-PA), an intimately related enzyme. Analysis of the selected peptide substrates strongly suggested that the primary sequence SGRSA, from position P2 to P2′, represents optimal subsite occupancy for substrates of u-PA. Insights gained in these investigations were used to design a variant of plasminogen activator inhibitor type 1, the primary physiological inhibitor of both u-PA and t-PA, that inhibited u-PA approximately 70 times more rapidly than it inhibited t-PA. These observations provide a solid foundation for the design of highly selective, high affinity inhibitors of u-PA and, consequently, may facilitate the development of novel therapeutic agents to inhibit the initiation and/or progression of selected human tumors.

Local activation and aggregation of platelets, followed by initiation of the blood coagulation cascade, assure that a fibrin clot will form rapidly in response to vascular injury (1). The presence of this thrombus, however, must be transient if the damaged tissue is to be remodeled and normal blood flow restored. The fibrinolytic system, which accomplishes the enzymatic degradation of fibrin, is therefore an essential component of the hemostatic system (1). The ultimate product of the fibrinolytic system is plasmin, a chymotrypsin family enzyme with relatively broad, trypsin-like primary specificity that is directly responsible for the efficient degradation of a fibrin clot (2). Production of this mature proteolytic enzyme from the inactive precursor, or zymogen, plasminogen is the rate-limiting step in the fibrinolytic cascade (2, 3). Catalysis of this key, regulatory reaction is tightly controlled *in vivo* and is mediated by two enzymes present in human plasma, u-PA[1] and t-PA (3–6).

u-PA and t-PA are very closely related members of the chymotrypsin gene family. These two proteases possess extremely high structural similarity (7, 8), share the same primary physiological substrate (plasminogen) and inhibitor (plasminogen activator inhibitor, type 1) (3), and, unlike plasmin, exhibit remarkably stringent substrate specificity (9–11). Despite their striking similarities, the physiological roles of t-PA and u-PA are distinct (5, 6), and many studies (5, 6, 12–18) suggest selective inhibition of either enzyme might have beneficial therapeutic effects. Mice lacking t-PA, for example, are resistant to specific excitotoxins that cause extensive neurodegeneration in wild type mice (13), and mice lacking u-PA exhibit defects in the proliferation and/or migration of smooth muscle cells in a model of restenosis following vascular injury (5, 6).

A large body of experimental evidence from studies involving both model systems and human patients suggests that u-PA may play an important role in tumor biology and provides a compelling rationale to pursue the development of u-PA inhibitors. For example, anti-u-PA antibodies inhibit metastasis of HEp3 human carcinoma cells to chick embryo lymph nodes, heart, and lung (19), and similar studies demonstrated that these antibodies inhibit lung metastasis in mice following injection of B16 melanoma cells into the tail vein (20). Anti-u-PA antibodies also inhibit both local invasiveness and lung metastasis in nude mice bearing subcutaneous MDA-MB-231 breast carcinoma tumors (21). In addition, a recent study indicated that u-PA-deficient mice are resistant to the induction and/or progression of several tumor types in a two-stage, chemical carcinogenesis model (18). Finally, high levels of tumor-associated u-PA correlate strongly with both a shortened disease-free interval and poor survival in several different human cancers (22–24).

Because mice lacking either u-PA or t-PA do not develop thrombotic disorders, selective inhibition of either of these two enzymes seems unlikely to create thrombotic complications *in vivo*. On the other hand, mice lacking both u-PA and t-PA suffer severe thrombosis in many organs and tissues, resulting in a significantly reduced life expectancy (5, 6). Nonselective inhibition of these two enzymes, therefore, seems almost certain to produce catastrophic consequences in the clinical setting. Consequently, significant interest exists in the development of inhibitors that are stringently specific for either t-PA

* This study was supported in part by National Institutes of Health (NIH) Grants RO1 HL52475 and PO1 HL31950 (to E. L. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ Supported by NIH Grant T32GMO8203.

‖ An assistant investigator with the Howard Hughes Medical Institute. To whom correspondence may be addressed: Dept. of Pharmacology and Howard Hughes Medical Institute, University of Texas Southwestern Medical Ctr., Dallas, TX 75235.

** To whom correspondence may be addressed: Dept. of Vascular Biology (VB-1), The Scripps Research Institute, 10550 N. Torrey Pines Rd., La Jolla, CA 92037. Fax: 619-784-7323.

[1] The abbreviations used are: u-PA, urokinase type plasminogen activator; t-PA, tissue type plasminogen activator; fTC, fAFF1-tether C; HPLC, high pressure liquid chromatography; PAI-1, plasminogen activator inhibitor type 1.

or u-PA, which are expected to facilitate a detailed investigation of the precise roles of the two enzymes in several important pathological processes and may aid the development of novel therapeutic agents to combat these processes. Rational design of these selective inhibitors is greatly complicated, however, by the absence of obvious "lead compounds"; both their primary physiological substrate and inhibitors fail to discriminate between the two closely related proteases.

We have used substrate phage display (25, 26) to elucidate optimal subsite occupancy of u-PA. Peptide substrates that match the consensus sequence for substrates of u-PA derived from these studies are cleaved by u-PA 840–5300 times more efficiently than control peptides containing the physiological target sequence present in plasminogen. In addition, unlike the plasminogen-derived control peptides, the selected peptides exhibit substantial selectivity for cleavage by u-PA versus t-PA. Information gained in these investigations was used to augment the u-PA/t-PA selectivity of PAI-1, the physiological inhibitor of both t-PA and u-PA (27, 28); suggests potential lead compounds for the design of selective, small molecule inhibitors of u-PA; and provides new insights into the divergent evolution of molecular recognition by intimately related enzymes.

MATERIALS AND METHODS

Reagents—Competent MC1061 (F) *Escherichia coli* and nitrocellulose were purchased from Bio-Rad Laboratories. Pansorbin (Protein A-bearing *Staphylococcus aureus*) cells were obtained from Calbiochem. K91 (F$^+$) and MC1061 (F$^+$) strains of *E. coli* were provided by Steve Cwirla (Affymax). mAb 3-E7 was purchased from Gramsch Laboratories (Schwabhausen, FRG). u-PA was obtained from Jack Henkin (Abbott Laboratories).

Construction of the phage vector fAFF1-tether C (fTC) and the random hexapeptide library fAFF-TC-LIB has been previously described (26). Control substrate phage fTC-P1$_5$ which contained the physiological target sequence for u-PA and t-PA, was constructed by hybridizing the single-stranded oligonucleotides 5′-TCGAGCGGTGGATCCGGTA-CTGGTCGTACTGGTCATGCTCTGGTAC-3′ and 5′-CGCCACCTAGG-CCAGGACCAGCACAACAACCACGAGAC-3′ and then ligating the annealed, double-stranded products into the *XhoI/Kpn*I-cut vector fTC. All constructs were first transformed into MC1061 by electroporation and then transferred into K91.

Measurement of Enzyme Concentrations—Concentrations of functional t-PA and u-PA were measured by active site titration with 4-methylumbelliferyl *p*-guanidinobenzoate (29) using a Perkin-Elmer LS 50B luminescence Fluorometer as described previously (9, 30). In addition, the enzymes were titrated with a standard PAI-1 preparation that had been previously titrated against a trypsin primary standard. Total enzyme concentrations were measured by enzyme-linked immunosorbent assay.

Phage Selection Using u-PA—Substrate phage display was originally developed by Matthews and Wells (25) using monovalent phage, and an alternative method that used multivalent phage was reported later by Smith and Navre (26). Multivalent substrate phage were screened with u-PA using reaction conditions identical to those previously reported for t-PA (31) except that digestion of the phage was performed using enzyme concentrations varying from 2 to 10 μg/ml and incubation times varying from 0.5 to 10 h.

Dot Blot Assay of Phage Proteolysis—Phage precipitation and dot blot analysis were performed as described previously (26, 31). Individual phage stocks were prepared and digested with no enzyme, t-PA, u-PA, or u-PA in the presence of 1 mM amiloride, a specific inhibitor of u-PA, for periods of time varying from 15 min to 10 h. Individual reaction mixtures were spotted onto a nitrocellulose filter using a dot blotter apparatus (Bio-Rad). The filter was probed with mAb 3E-7 and developed using the Amersham Western ECL kit. Loss of positive staining indicates loss of antibody epitopes from the phage due to proteolytic cleavage of the randomized hexamer region.

Preparation and Sequencing of DNA from Phage Clones—DNA samples were prepared from interesting phage clones as described previously (31). Briefly, phage were precipitated from a 1-ml overnight culture by adding 200 μl of 20% polyethylene glycol in 2.5 M NaCl. The mixture was incubated on ice for 30 min, and the phage pellet was collected by microcentrifugation for 5 min. The phage were resuspended in 40 μl of lysis buffer (10 mM Tris-HCl, pH 7.6, 0.1 mM EDTA, 0.5% Triton X-100) and heated at 80 °C for 15 min. Single-stranded DNA was purified by phenol extraction and ethanol precipitation and sequenced by the dideoxy method.

Kinetics of Cleavage of Synthetic Peptides by t-PA and u-PA—Peptides were synthesized and purified as described (9). Kinetic data were obtained by incubating various concentrations of peptide with a constant enzyme concentration to achieve between 5 and 20% cleavage of the peptide in each reaction. For assays with u-PA, enzyme concentration was either 815 or 635 nM. For assays with t-PA enzyme, concentration was 700 nM. Peptide concentrations were chosen where possible to surround K_m and in all cases were between 0.5 and 32 mM. The buffer used in these assays has been described (9). Reactions were stopped by the addition of trifluoroacetic acid to 0.33% or by freezing on dry ice. Cleavage of the 13- and 14-residue peptides was monitored by reverse phase HPLC as described (9). The 4–6-residue peptides were acylated at their amino termini and amidated at their carboxyl termini. Cleavage of the 4–6-residue peptides was monitored by hydrophilic interaction HPLC chromatography (32) using a polyhydroxyaspartamine column from PolyLC (Columbia, MD). Buffer A was 50 mM triethylamine phosphate in 10% acetonitrile, and buffer B was 10 mM triethylamine phosphate in 80% acetonitrile. Peptides were eluted by a gradient that was varied from 100% buffer B to 100% buffer A during a 13-min interval. The percentage of cleaved peptide was calculated by dividing the area under the product peaks by the total area under substrate and product peaks. For all peptides containing multiple basic residues, mass spectral analysis of products confirmed that cleavage occurred at a single site and identified the scissile bond. Data were interpreted by Eadie-Hofstee analysis. Errors were determined as described (33) and were <25%.

Site-directed Mutagenesis and Construction of an Expression Vector Encoding a Recombinant Variant of PAI-1—The expression vector pPAIST7HS was derived from the plasmid pBR322 and contained a full-length cDNA encoding human PAI-1 that was transcribed from a T7 gene 10 promoter (34). The 300-base pair *SalI/BamH*I fragment of human PAI-1 was subcloned from pPAIST7HS into bacteriophage M13mp18. Single-stranded DNA produced by the recombinant M13mp18 constructs was used as a template for site-specific mutagenesis according to the method of Zoller and Smith (35) as modified by Kunkel (36). The mutagenic oligonucleotide had the sequence 5′-CCACAGCTGTCATAGGCAGCGGCAAAAGCGCCCCCGAGGAGA-TC-3′.

Following mutagenesis, single-stranded DNA corresponding to the entire 300-base pair *SalI-BamH*I fragment was fully sequenced to ensure the presence of the desired mutations and the absence of any additional mutation. The 300-base pair *SalI-BamH*I double-stranded DNA fragment from the mutated, replicative form DNA was used to replace the corresponding fragment in pPAIST7HS to yield a full-length cDNA encoding PAI-1/UK1, which contained the amino acid sequence GGSRSA from the P4 to P2′ position of the reactive center loop.

Expression and Purification of Recombinant Wild Type PAI-1 and the Variant PAI-1/UK1—Expression of wild type and the mutated variant of PAI-1 was accomplished in the *E. coli* strain BL21[DE3]pLys^s (Novagen), which synthesizes T7 RNA polymerase in the presence of isopropyl-1-thio-β-D-galactopyranoside. Bacterial cultures were grown at 37 °C with vigorous shaking to an A_{595} of 0.9–1.1, and isopropyl-1-thio-β-D-galactopyranoside was added to a final concentration of 1 mM to induce the synthesis of T7 RNA polymerase and the production of PAI-1 proteins. Cultures were grown for an additional 1–2 h at 37 °C and then shifted to 30 °C for 2–6 h.

Cells were pelleted by centrifugation at 8000 × *g* for 20 min at 4 °C and resuspended in 40 ml of cold start buffer (20 mM sodium acetate, 200 mM NaCl and 0.01% Tween 20, pH 5.6). The cell suspension was disrupted in a French pressure cell (Aminco), and cellular debris was removed by ultracentrifugation for 25 min at 32,000 × *g*.

Purification of soluble, active PAI-1 was performed as described previously (37). PAI-1 containing supernatants were injected onto a XK-26 column (Pharmacia Biotech Inc.) packed with CM-50 Sephadex (Pharmacia). The column was washed with 5 column volumes of start buffer (20 mM sodium acetate, 200 mM NaCl, and 0.01% Tween 20, pH 5.6), and PAI-1 proteins were eluted using a 0.2–1.8 M linear gradient of NaCl in the same buffer. Peak fractions were collected, pooled, and concentrated using a Centriplus 30 concentrator (Amicon). Purified preparations were analyzed by activity measurements using standard, direct assays of t-PA, SDS-polyacrylamide gel electrophoresis, and measurement of optical density at 280 nm.

Measurement of Active PAI-1 in Purified Preparations—A primary standard of trypsin was prepared by active site titration using *p*-nitrophenyl guanidinobenzoate HCl as described previously (38). Concentra-

TABLE I
Amino acid sequences of the randomized hexapeptide in 91 isolated substrate phage clones

Clone number	Amino acid sequence	Clone number	Amino acid sequence
1	S G R A R Q	47	(S G)R K V P G S
2	S K S G R S(L)	48	(S G)R K N I S G
3	S S R N A D	49	(S G)R L A T K A
4	T A R L R G	50	(S G)R M R K N D
5	T A R S D N	51	(S G)R N A Q V R
6	T S R M G T	52	(S G)R N A V E P
7	T S R Q A Q	53	(S G)R N D R L N
8	T T R R N K	54	(S G)R N G K S R
9	T T S R R S	55	(S G)R N M P L L
10	W S G R S G	56	(S G)R N T G S H
11	A I K R S A	57	(S G)R R M T M G
12	(G)G R R G N R	58	(S G)R R R L N M
13	(G)G R S V N N	59	(S G)R R T L D F
14	H T R R M K	60	(S G)R R A V S N
15	I S T A R M(L)	61	(S G)R S A K V D
16	(S G)K A A D V T	62	(S G)R S A N A I
17	K K R T N D	63	(S G)R S A T R D
18	K M S A R I(L)	64	(S G)R S A V V K
19	(G)K R R D V A	65	(S G)R S D Q F L
20	(G)K R V S K N	66	(S G)R S D N P N
21	(S G)K S A D A A	67	(S G)R S E R S L
22	(S G)R A A A M V	68	(S G)R S G D P G
23	(S G)R A G N I R	69	(S G)R S G N T T
24	(S G)R A H R D N	70	(S G)R S G N M G
25	(S G)R A R D D R	71	(S G)R S H N V G
26	(S G)R A R H M V	72	(S G)R S P D G M
27	(S G)R A R S P R	73	(S G)R S R R L P
28	(S G)R A V G H Q	74	(S G)R S R V T S
29	(S G)R A V V D S	75	(S G)R S S H S S
30	(S G)R G G K G P	76	(S G)R S S O A A
31	(S G)R G R S A V	77	(S G)R S S S S H
32	(S G)R G V D M N	78	(S G)R S S S T V
33	(S G)R G V K M H	79	(S G)R S T D L V
34	(S G)R H R S D I	80	(S G)R S T N V E
35	(S G)R K G Q G G	81	(S G)R S T R H K
36	(S G)R K L H M N	82	(S G)R S Y T N S
37	(S G)R K M D M G	83	(S G)R T S P S T
38	(S G)R K M D R S	84	(S G)R T S V N L
39	(S G)R K M R M G	85	S K R A S I
40	(S G)R K N Q R V	86	S Q T C V R(L V)
41	(S G)R K Q R D S	87	T E R R V R(L V)
42	(S G)R K R V G A	88	T Q R S T G
43	(S G)R K S K V V	89	T R R D R I
44	(S G)R K S T S S	90	V A R N Y K
45	(S G)R K V G S L	91	V S R R N M
46	(S G)R K A S L S		

tions of active molecules in purified preparations of wild type or mutated PAI-1s were determined by titration of standardized trypsin as described by Olson *et al.* (39) and by titration of standardized t-PA preparations.

Kinetic Analysis of the Inhibition of t-PA and u-PA by Recombinant PAI-1 and PAI-1/UK1—Second order rate constants (k_i) for inhibition of t-PA or u-PA were determined using pseudo-first order ($k_i < 2 \times 10^6$) or second order ($k_i > 2 \times 10^6$) conditions. For each reaction, the concentrations of enzyme and inhibitor were chosen to yield several data points for which the residual enzymatic activity varied between 20 and 80% of the initial activity. Reaction conditions and data analysis for pseudo-first order reactions were as described previously (40–43).

For second order reactions, equimolar concentrations of u-PA and PAI-1 were mixed directly in microtiter plate wells and preincubated at room temperature for periods of time varying from 0 to 30 min. Following preincubation, the mixtures were quenched with an excess of neutralizing anti-PAI-1 antibody (generously provided by Dr. David Loskutoff), and residual enzymatic activity was measured using a standard, indirect chromogenic assay. These indirect, chromogenic assays were compared with control reactions containing no PAI-1 or to which PAI-1 was added after preincubation and the addition of anti-PAI-1 antibody, plasminogen, and Spec PL to the reaction mixture. Data were analyzed by plotting the reciprocal of the residual enzyme concentration *versus* the time of preincubation.

RESULTS

Construction and Use of Substrate Phage Libraries—A polyvalent fd phage library that displayed random hexapeptide sequences and contained 2×10^8 independent recombinants was prepared (25, 31). Each member of this library displayed an N-terminal extension from phage coat protein III that contained a randomized region of six amino acids, a six-residue linker sequence (SSGGSG), and the epitopes for mAbs 179 and 3-E7. Because u-PA did not digest the phage coat protein III sequence, the antibody epitopes, or the flexible linker sequence, the loss of antibody epitopes from the phage surface upon incubation with u-PA required cleavage of the randomized peptide insert. Incubation of the library with u-PA, followed by removal of phage retaining the antibody epitopes, therefore, accomplished a large enrichment of phage clones whose random hexamer sequences could be cleaved by u-PA.

Analysis of Selected Phage Clones and Identification of a Consensus Sequence—Following five rounds of selection to enrich and amplify phage that display sequences that are readily cleaved by u-PA, 100 phage clones were identified as u-PA substrates. DNA sequencing of these clones revealed the presence of 91 distinct hexamer sequences among the selected phage (Table I). As expected from the trypsin-like primary specificity of u-PA, each hexamer contained at least one basic residue, and 89 of the 91 hexamer sequences contained at least one arginine residue. 35 of the 91 substrate phage contained a single basic residue, and in 33 of these 35 cases the single basic

TABLE II
Amino acid sequence of the randomized hexamer present in the most labile u-PA substrate phage

Clone number	P3	P2	P1	P1′	P2′	P3′	P4′	P5′
46	(S	G)	R	K	A	S	L	S
51	(S	G)	R	N	A	Q	V	R
60	(S	G)	R	R	A	V	S	N
61	(S	G)	R	S	A	K	V	D
62	(S	G)	R	S	A	N	A	I
63	(S	G)	R	S	A	T	R	D
64	(S	G)	R	S	A	V	V	K
77	(S	G)	R	S	S	S	S	H

residue was an arginine. An additional 22 phage contained two basic residues but only a single arginine. Alignment and analysis of these hexamer sequences suggested that the consensus sequence for optimal subsite occupancy for substrates of u-PA, from P3 to P2′, was SGR(S > R,K,A)*X*, where *X* represents a variety of amino acid residues but was most often alanine, glycine, serine, valine, or arginine.

Analysis of these data was complicated by the fact that approximately 72% of the selected substrate phage contained an arginine in the first position of the randomized hexamer and therefore utilized the amino-terminal flanking residues, Ser-Gly, to occupy the P3 and P2 subsites. While these results left no doubt that the P3–P1 SGR sequence created by the fusion was a very favorable recognition site for u-PA, this use of flanking residues necessitated a particularly careful examination of the P3 and P2 preferences of u-PA. Consequently, we altered our experimental protocol in two ways to address this issue. First, we isolated an unusually large collection of substrate phage (91 distinct substrates) to ensure that a reasonable number of these (23) would not utilize the flanking Ser-Gly to fill the P3 and P2 subsites. This allowed a meaningful comparison of the consensus sequence derived from the entire library with that derived from the non-fusion phage and the demonstration of good agreement between the two consensus sequences. Second, we performed a previously described dot blot analysis (26, 31) of the digestion of all 100 substrate phage by u-PA using a wide variety of stringencies of digestion. Although this semiquantitative assay cannot provide kinetic constants, it can provide an accurate rank ordering of the lability of the substrate phage clones.

Under the most stringent conditions examined, 11 of the 100 substrate phage, containing eight distinct randomized hexamer sequences, proved to be particularly labile u-PA substrates (Table II). All eight of the most labile substrate phage contained the P3–P1 SGR motif, demonstrating that this sequence is, in fact, a more labile u-PA site than related, selected sequences present in the library such as SSR, TAR, TSR, TTR, etc. This dot blot analysis also yielded additional information regarding the preferences of u-PA for the unprimed subsites. While analysis of the entire substrate phage library failed to reveal a clear consensus at P1′ and P2′, the most labile substrate phage displayed an obvious preference at both of these positions. Five of the eight most labile phage contained a serine residue at P1′, and seven of these eight phage contained an alanine residue at P2′. These observations strongly suggest that the primary sequence SGRSA, from P3 to P2′, represents optimal subsite occupancy for substrates of u-PA.

Kinetic Analysis of the Cleavage of Peptides Containing Sequences Present in Selected Substrate Phage—Four peptides containing amino acid sequences present in the randomized hexamer region of the most labile phage were chosen for detailed kinetic analysis (Table III) and compared with the hydrolysis of a control peptide (I) containing the P3–P4′ sequence of plasminogen, a series of residues that fall within a disulfide-linked loop in the native protein. All four of the selected peptides were substantially improved substrates for u-PA, by factors of 840–5300, compared with the control, plasminogen peptide (Table III). These increases in catalytic efficiency were mediated primarily by increases in k_{cat}, suggesting that optimized subsite interactions served to lower the energy of the transition state rather than the ground state. For example, compared with that of control peptide (I), the K_m for cleavage of the most labile selected peptide (II) was reduced by a factor of 5.6; however, the k_{cat} was increased by a factor of more than 940. In addition, peptide substrates that interacted optimally with the primary subsites of u-PA were selective for cleavage by u-PA relative to t-PA. The four selected peptides (II–V), for example, were cleaved 16–89 times more efficiently by u-PA than by t-PA, and improvements in both K_m and k_{cat} contributed to the preferential hydrolysis by u-PA.

Minimization of the Selective Peptide Substrates—The kinetic analysis described above was performed using substrate peptides that were 14 amino acids in length. To confirm that the specificity we observed was inherent in the selected hexapeptide sequences, we examined the kinetics of cleavage of short peptides containing only sequences found within selected hexapeptide sequences. Pentapeptide VII, for example, was cleaved by u-PA with a catalytic efficiency of 1200 M^{-1} s^{-1} and exhibited a u-PA/t-PA selectivity of 20. The behavior of pentamer VII in these assays, therefore, was very similar to that of peptide IV, a 14-mer that contains the same P3–P2′ sequence as the pentamer. These observations indicate that appropriate occupancy of the P3–P2′ subsites alone can create selective substrates for u-PA.

Effect of Lysine versus Arginine at P1—Differences at position 190 (chymotrypsin numbering system) between u-PA and t-PA suggest that u-PA may exhibit decreased discrimination between arginine and lysine at the P1 position of a substrate compared with t-PA (44). Consistent with this hypothesis and in contrast to the selected t-PA substrate library, the u-PA library did include members that contained a P1 lysine. This observation suggested that the u-PA/t-PA selectivity of a peptide substrate might be enhanced by placement of lysine in the P1 position, although this increased selectivity was likely to be accompanied by decreased reactivity toward u-PA. To test this hypothesis, we analyzed hydrolysis of a variant of u-PA selective peptide (VI) that contained a P1 lysine (peptide VIII). The P1 lysine mutation decreased the catalytic efficiency for cleavage of this peptide by a factor of 49 for t-PA and by a factor of 7 for u-PA. As predicted, then, the P1 lysine mutation did enhance the u-PA/t-PA selectivity of the peptide substrate by a factor of approximately 7. It is not surprising, therefore, that the most selective u-PA substrate, peptide IX, which is cleaved approximately 121 times more efficiently by u-PA than by t-PA, is derived from the randomized hexamer region of a substrate phage that contained a P1 lysine.

Importance of P3 and P4 for Discrimination between u-PA and t-PA—Recent investigations that explored optimal subsite occupancy for substrates of t-PA suggested that the P3 residue was the primary determinant of the ability of a substrate to discriminate between t-PA and u-PA and that this selectivity could be enhanced modestly by appropriate occupancy of P4 (11). These suggestions were based on evidence obtained from a statistical analysis of phage selected using a substrate subtraction protocol rather than by a kinetic analysis of peptide substrates. Consequently, to test these hypotheses, we synthesized variants of the most labile u-PA-selective substrate (peptide II) that contained mutations in the P3 and/or P4 positions and analyzed the hydrolysis of these peptides by u-PA and t-PA. In peptide X the P3 serine of peptide II was replaced by

TABLE III

Comparison of k_{cat}, k_m, and k_{cat}/K_m for the hydrolysis by t-PA or u-PA of peptides selected for preferential cleavage by u-PA

Substrate (Pn . . . P3, P2, P1 ↓ P1′, P2′, P3′ . . . Pn′)*	u-PA			t-PA			u-PA:t-PA selectivity
	k_{cat}	K_m	k_{cat}/K_m	k_{cat}	K_m	k_{cat}/K_m	
	s^{-1}	μM	$M^{-1}\ s^{-1}$	s^{-1}	μM	$M^{-1}\ s^{-1}$	
Native cleavage sequence from plasminogen							
I KKSPGR ↓ VVGGSVAH	0.003	3400	0.88	0.0043	15000	0.29	3.0
u-PA-selective peptides							
II LGGSGR ↓ SANAILE	2.83	603	4700	0.305	4080	75	63
III LGGSGR ↓ NAQVRLE	3.69	1160	3200	0.255	7000	36	89
IV LGGSGR ↓ SATRDLE	0.54	733	740	0.068	1500	45	16
V LGGSGR ↓ KASLSLE	1.14	1130	1010	0.168	5100	33	31
Minimized, u-PA-selective peptides							
VI SGR ↓ S	2.3	2100	1100	5.0	15,000	330	3.3
VII SGR ↓ SA	3.7	3100	1200	2.4	40,000	60	20
VIII SGK ↓ S	1.22	7900	154	0.19	28,000	6.8	23
IX GSGK ↓ S	0.82	4250	193	0.07	44,000	1.6	121
Variants of u-PA-selective peptides							
X LGGYGR ↓ SANAILE	0.7	2200	318	3.29	1850	1800	0.018
XI LGGRGR ↓ SANAILE	0.08	1200	67	0.85	2400	350	0.019
XII LGQRGR ↓ SANAILE	0.068	1500	45	2.55	3000	850	0.005

* Positional nomenclature of subsite residues. Arrows denote the position of peptide bond hydrolysis. The peptide bond is cleaved between P1 and P1′. The error in these determinations was 4–22%.

a tyrosine, and in peptide XI the P3 serine was replaced by arginine. As expected, these mutations substantially decreased the u-PA/t-PA selectivity of the peptide by a factor of 330 or 360, respectively, and actually converted the peptide into a t-PA-selective substrate. Moreover, mutation of both the P3 serine and P4 glycine of the most labile u-PA substrate to arginine and glutamine, respectively (peptide XII), decreased the u-PA/t-PA selectivity by a factor of 1200. These data confirm the proposed status of the P3 and P4 residues as specificity determinants for substrates of t-PA and u-PA and suggest a particularly prominent role of the P3 residue in this capacity.

Design and Characterization of a Variant of PAI-1 That Is Selective for u-PA—To test the prediction that information gained from the study of peptide substrates could facilitate the design of selective, high affinity inhibitors of urokinase, we sought to augment the u-PA/t-PA selectivity of the serpin PAI-1, the primary physiological inhibitor of both t-PA and u-PA. We used oligonucleotide-directed, site-specific mutagenesis to construct a variant of PAI-1 that contained the primary sequence found in the peptide substrate that was most selective for u-PA, GSGKS, from the P4–P1′ position of the reactive center loop. Kinetic analysis indicated that the PAI-1 variant inhibited u-PA approximately 70 times more rapidly than it inhibited t-PA with second order rate constants for inhibition of u-PA and t-PA of $6.2 \times 10^6\ M^{-1}\ s^{-1}$ and $9 \times 10^4\ M^{-1}\ s^{-1}$, respectively (Table IV). In contrast, wild type PAI-1 inhibits u-PA and t-PA with second order rate constants of $1.9 \times 10^7\ M^{-1}\ s^{-1}$ and $1.8 \times 10^6\ M^{-1}\ s^{-1}$, respectively. As anticipated, therefore, the mutated serpin possessed a u-PA/t-PA selectivity that was approximately 7-fold greater than that of wild type PAI-1. Moreover, the 70-fold selectivity of the PAI-1 variant is consistent with the value of 120 observed for hydrolysis of the corresponding peptide substrate by the two enzymes (Tables III and IV).

DISCUSSION

Substrate Phage Can Elucidate Specificity Differences between Closely Related Enzymes—u-PA and t-PA possess distinct but overlapping physiological and pathological roles, and the ability to selectively inhibit either enzyme with small molecules would allow these roles to be examined comprehensively both *in vitro* and *in vivo*. Normally, such inhibitor design would be based on knowledge of the sequences of endogenous protein substrates or inhibitors. This approach, however, is not possible with u-PA and t-PA because these enzymes share the same physiological substrate, plasminogen, and inhibitor, PAI-1. Furthermore, this similarity calls into question the hypothesis that highly selective inhibitors can be generated, since the specificities of the two enzymes appear so similar. We find, however, both in this study and in a previous study aimed at the design of t-PA selective substrates (11), that there are subtle but significant differences in optimal subsite occupancy between the two enzymes, and these distinctions can be elucidated by substrate phage display protocols.

Sequences Selected for Optimal Cleavage Do Not Resemble the Physiological Target Sequence—A key observation of this study is that the primary sequence SGRSA, from the P3–P2′ positions of a peptide substrate, affords highly labile subsite occupancy for urokinase. This sequence differs at P3, P1′, and P2′ from the target sequence found in plasminogen (PGRVV) and is cleaved by u-PA greater than 5300 times more efficiently. This major discrepancy, in both primary sequence and lability, of the physiological target sequence and the consensus sequence derived using substrate phage display protocols suggests that a physiological target sequence is not necessarily a reasonable lead compound for the design of specific, small molecule substrates or inhibitors of highly selective serine proteases.

A major contribution to the discrepancy between the physiological and consensus target sequences of u-PA almost certainly arises from the highly conserved mechanism of zymogen activation of chymotrypsin family enzymes (45). Following activation cleavage of a chymotrypsinogen-like zymogen, the P1′ and P2′ residues insert into the activation pocket, where they form a number of conserved hydrophobic interactions as well as a new, buried salt bridge with the aspartic acid residue adjacent to the active site serine (45–47). Because these interactions substantially stabilize the active conformation of the mature enzyme, this key role after activation cleavage places severe functional constraints on the P1′ and P2′ residues of a chymotrypsinogen-like zymogen and consequently prevents the two residues from evolving simply to interact optimally with the activating enzyme. Consistent with this hypothesis, the consensus and physiological target sequences for u-PA agree well on the unprimed side of the scissile bond; however, the two target sequences diverge dramatically at the P1′ and P2′ subsites.

Additional factors are also likely to contribute to the observed discrepancy between the consensus and physiological

TABLE IV
Second order rate constants for inhibition of t-PA or u-PA by wild type PAI-1 and PAI-1/UK1

Inhibitor	Primary sequence of reactive center loop (P4–P2')	Rate constant toward u-PA	Rate constant toward t-PA	u-PA/t-PA selectivity
		$M^{-1}\,s^{-1}$	$M^{-1}\,s^{-1}$	
Wild type PAI-1	VSAR ↓ MA	1.9×10^7	1.8×10^6	11
PAI-1/UK1	GSGK ↓ SA	6.2×10^6	9.0×10^4	69

target sequences for u-PA. For example, modeling studies reported by Lamba, Huber, Bode and co-workers (8) suggest the S1' and/or the S2' pockets utilized by u-PA when hydrolyzing plasminogen may actually differ from those used when hydrolyzing peptide substrates. Moreover, as the enzyme diverged from a trypsin-like precursor, u-PA may have evolved a strong dependence for efficient catalysis upon productive interactions with substrates at secondary sites that diminished the contribution of optimal interactions with primary subsites in the active site cleft. Although the location, role, and even the existence of such secondary contacts between u-PA and plasminogen remain obscure at the present time, previous studies of the interaction of u-PA and t-PA with PAI-1 have demonstrated very clearly that these two enzymes are capable of using specific, secondary contacts efficiently both to enhance selectivity and to dampen the influence of optimal primary subsite interactions (42, 48–50). Although the reactive center loop of PAI-1 has evolved to match optimal subsite occupancy for urokinase very closely, in the absence of productive contact with a single, strong secondary site of interaction between the two proteins, PAI-1 becomes a poor inhibitor of u-PA (50).

Implications Regarding the Possibility of Additional Physiological Substrates for u-PA—The identification of synthetic peptides that are cleaved up to 120 times more efficiently by u-PA than by t-PA raises the possibility that similar u-PA-selective (or t-PA-selective) physiological substrates may exist that are currently not appreciated. Differences in the phenotypes exhibited by mice lacking either of the two enzymes are consistent with this possibility (5, 6). This issue remains uncertain, however, because selective expression of t-PA or u-PA in particular microenvironments could also account for these distinct phenotypes.

Importance of the P3 Residue in Discriminating between u-PA and t-PA—By demonstrating that mutation of the P3 residue alone could alter the relative u-PA/t-PA selectivity of a peptide substrate by a factor of greater than 300 (Table III), this study provided strong support for the hypothesis that the P3 residue was the primary determinant of the ability of a substrate to discriminate between u-PA and t-PA. We have previously reported that occupancy of P3 by arginine or large aromatic or hydrophobic residues favored cleavage by t-PA (11), and this investigation showed that a P3 serine residue favored cleavage by u-PA. In addition, this study demonstrated that more modest alterations of specificity could be achieved by selective occupancy of the P4 and P1 subsites. These data indicated that PAI-1, which contains a P3 serine, has evolved to match optimal subsite occupancy of u-PA more closely than that of t-PA. This observation may explain why PAI-1 inhibits u-PA more rapidly than it inhibits t-PA (Table IV) and suggests that, during the evolution of the fibrinolytic system, there may have been a greater need to suppress the activity of u-PA in the circulation than to regulate t-PA activity. Consistent with this hypothesis, the circulating, single chain form of u-PA is a true zymogen, while t-PA is secreted into the circulation as an active, single chain enzyme.

Substrate Phage Display Can Aid Inhibitor Design—Another implication of these studies is that information gained from the application of substrate phage display libraries can lead directly to the design of specific inhibitors. Although hydrolysis of the selective, small peptide substrates by u-PA is characterized by K_m values in the 0.6–3 mM range, it has been routinely observed that the introduction of a transition state bond geometry adjacent to the P1 residue of a protease substrate can create either a reversible inhibitor whose affinity for the target protease is enhanced by 3–6 orders of magnitude or an irreversible inhibitor with an impressive second order rate constant for inhibition of the target protease ($>10^8\ M^{-1}\,s^{-1}$) (for a review, see Ref. 51). Similar results using the substrates identified in this study would create highly selective, small molecule u-PA inhibitors, with affinities in the low nanomolar range, that might be further improved by subsequent, systematic chemical modification.

Conclusion—The ability to identify subtle but significant specificity differences between enzymes that share the same physiological substrates and inhibitors, as demonstrated in this study, is a fundamental challenge both for basic enzymology and rational drug design. Advances in this area will significantly enhance understanding of the molecular determinants and mechanisms of specific catalysis and may facilitate the design of highly selective and therapeutically valuable new enzymes.

Acknowledgments—We thank Steven Madden, Bikash C. Pramanik, and Dr. Clive Slaughter for peptide synthesis and mass spectral analysis.

REFERENCES

1. Roberts, H. R., and Tabares, A. H. (1995) in *Molecular Basis of Thrombosis and Hemostasis* (High, K. A., Roberts, H. R., eds) pp. 35–50, Marcel Dekker, New York
2. Castellino, F. J. (1995) in *Molecular Basis of Thrombosis and Hemostasis* (High, K. A., Roberts, H. R., eds) pp. 495–515, Marcel Dekker, New York
3. Collen, D., and Lijnen, H. R. (1991) *Blood* **78**, 3114–3124
4. Madison, E. L. (1994) *Fibrinolysis* **8**, Suppl. 1, 221–236
5. Carmeliet, P., Schoonjans, L., Kieckens, L., Ream, B., Degen, J., Bronson, R., De Vos, R., van den Oord, J. J., Collen, D., and Mulligan, R. C. (1994) *Nature* **368**, 419–424
6. Carmeliet, P., and Collen, D. (1996) *Fibrinolysis* **10**, 195–213
7. Spraggon, G., Phillips, C., Nowak, U. K., Ponting, C. P., Saunders, D., Dobson, C. M., Stuart, D. I., and Jones, E. Y. (1995) *Structure* **3**, 681–691
8. Lamba, B., Bauer, M., Huber, R., Fischer, S., Rudolph, R., Kohnert, U., and Bode, W. (1996) *J. Mol. Biol.* **258**, 117–135
9. Madison, E. L., Coombs, G. S., and Corey, D. R. (1995) *J. Biol. Chem.* **270**, 7558–7562
10. Coombs, G. S., Dang, A. T., Madison, E. L., and Corey, D. R. (1996) *J. Biol. Chem.* **271**, 4461–4467
11. Ke, S.-H., Coombs, G. S., Tachias, K., Navre, M., Corey, D. R., and Madison, E. L. (1997) *J. Biol. Chem.* **272**, 16603–16609
12. Dano, K., Andreasen, P. A., Grondahl-Hansen, J., Kristensen, P., Nielsen, L. S., and Skriver, L. (1985) *Adv. Cancer Res.* **44**, 139–266
13. Tsirka, S. E., Gualandris, A., Amaral, D. G., and Strickland, S. (1995) *Nature* **377**, 340–344
14. Fukudome, K., and Esmon, C. T. (1994) *J. Biol. Chem.* **269**, 26486–26491
15. Min, H. Y., Doyle, L. V., Vitt, C. L., Zandonella, C. L., Stratton-Thomas, J. R., Shuman, M. A., and Rosenberg, S. (1996) *Cancer Res.* **56**, 2428–2433
16. Seeds, N. W., Williams, B. L., and Bickford, P. C. (1995) *Science* **270**, 1992–1994
17. Osozowski, L. (1988) *Cell* **52**, 321–328
18. Shapiro, R. L., Duquette, J. G., Roses, D. F., Nunes, I., Harris, M. N., Kamino, H., Wilson, E. L., and Rifkin, D. B. (1996) *Cancer Res.* **56**, 3597–3604
19. Osozowski, L., and Reich, E. (1983) *Cell* **35**, 611–619
20. Hearing, V. J., Law, L. W., Corti, A., Appella, E., and Blasi, F. (1988) *Cancer Res.* **48**, 1270–1276
21. Brünner, N., Heyer-Hansen, G., Remer, J., Ellis, V., Holst-Hansen, C., Spang-Thomsen, M., and Dana, K. (1992) *Proc. Am. Assoc. Cancer Res.* **33**, 61
22. Duffy, M. J., O'Grady, P., Devaney, D., O'Siorain, L., Fennelly, J. J., and Lijnen, H. J. (1988) *Cancer* **62**, 531–533
23. Jänicke, F., Schmitt, M., Hafter, R., Hollrieder, A., Bebic, R., Ulm, K., Gössner, W., and Graeff, H. (1990) *Fibrinolysis* **4**, 69–78

24. Duffy, M. J. (1993) *Fibrinolysis* **7**, 295–302
25. Matthews, D. J., and Wells, J. A. (1993) *Science* **260**, 1113–1117
26. Smith, M. M., Shi, L., and Navre, M. (1995) *J. Biol. Chem.* **270**, 6440–6449
27. Loskutoff, D. J. (1993) *J. Clin. Invest.* **92**, 2563
28. van Meijer, M., and Pannekoek, H. (1995) *Fibrinolysis* **9**, 263–276
29. Jameson, G. W., Roberts, D. V., Adams, R. W., Kyle, S. A., and Ellmore, D. T. (1973) *Biochem. J.* **131**, 107–117
30. Urano, T., Urano, S., and Castellino, F. J. (1988) *Biochem. Biophys. Res. Commun.* **156**, 45–51
31. Ding, L., Coombs, G. S., Strandberg, L., Navre, M., Corey, D. R., and Madison, E. L. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 7627–7631
32. Alpert, A. J. (1990) *J. Chromatogr.* **499**, 177–196
33. Taylor, J. R. (1982) *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*, University Science Books, Mill Valley, CA
34. Tucker, H. M., Mottonen, J., Goldsmith, E. J., and Gerard, R. D. (1995) *Nat. Struct. Biol.* **2**, 442–445
35. Zoller, M. J., and Smith, M. (1984) *DNA* **3**, 479–488
36. Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 488–492
37. Sancho, E., Tonge, D. W., Hockney, R. C., and Booth, N. A. (1994) *Eur. J. Biochem.* **224**, 126–134
38. Chase, T., and Shaw, E. (1967) *Biochem. Biophys. Res. Commun.* **29**, 508–514
39. Olson, S. T., Bock, P. E., Kvassman, J., Shore, J. D., Lawrence, D. A., Ginsburg, D., and Björk, I. (1995) *J. Biol. Chem.* **270**, 30007–30017
40. Holmes, W. E., Lijnen, H. R., and Collen, D. (1987) *Biochemistry* **26**, 5133–5140
41. Beatty, K., Beth, J., and Travis, J. (1980) *J. Biol. Chem.* **255**, 3931–3934
42. Madison, E. L., Goldsmith, E. J., Gerard, R. D., Gething, M. J., Sambrook, J. F., and Bassel-Duby, R. S. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 3530–3533
43. Madison, E. L., and Sambrook, J. F. (1993) *Methods Enzymol.* **223**, 249–271
44. Bode, W., Turk, D., and Karshikov, A. (1992) *Protein Sci.* **1**, 426–471
45. Huber, R., and Bode, W. (1978) *Acc. Chem. Res.* **11**, 114–122
46. Freer, S. T., Kraut, J., Robertus, J. D., Wright, H. T., and Xuong, N. H. (1970) *Biochemistry* **9**, 1997–2009
47. Fehlhammer, H., Bode, W., and Huber, R. (1977) *J. Mol. Biol.* **111**, 415–438
48. Madison, E. L., Goldsmith, E. J., Gerard, R. D., Gething, M.-J., and Sambrook, J. F. (1989) *Nature* **339**, 721–724
49. Madison, E. L., Goldsmith, E. J., Gething, M. J., Sambrook, J. F., and Gerard, R. D. (1990) *J. Biol. Chem.* **265**, 21423–21426
50. Adams, D. S., Griffin, L. A., Nachajko, W. R., Reddy, V. B., and Wei, C.-M. (1991) *J. Biol. Chem.* **266**, 8476–8482
51. Baggio, R., Shi, Y.-Q., Wu, Y.-Q., and Abeles, R. H. (1996) *Biochemistry* **35**, 3551–3553

THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 1997 by The American Society for Biochemistry and Molecular Biology, Inc.

Vol. 272, No. 26, Issue of June 27, pp. 16603–16609, 1997
Printed in U.S.A.

Distinguishing the Specificities of Closely Related Proteases

ROLE OF P3 IN SUBSTRATE AND INHIBITOR DISCRIMINATION BETWEEN TISSUE-TYPE PLASMINOGEN ACTIVATOR AND UROKINASE*

(Received for publication, November 19, 1996, and in revised form, April 16, 1997)

Song-Hua Ke‡, Gary S. Coombs§, Kathy Tachias‡, Marc Navre¶, David R. Corey§‖, and Edwin L. Madison‡

From the ‡Department of Vascular Biology, The Scripps Research Institute, La Jolla, California 92037, the §Department of Pharmacology and Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, Texas 75235, and ¶Affymax Research Institute, Santa Clara, California 95051

Elucidating subtle specificity differences between closely related enzymes is a fundamental challenge for both enzymology and drug design. We have addressed this issue for two intimately related serine proteases, tissue-type plasminogen activator (t-PA) and urokinase-type plasminogen activator (u-PA), by modifying the technique of substrate phage display to create substrate subtraction libraries. Characterization of individual members of the substrate subtraction library accomplished the rapid, direct identification of small, highly selective substrates for t-PA. Comparison of the amino acid sequences of these selective substrates with the consensus sequence for optimal substrates for t-PA, derived using standard substrate phage display protocols, suggested that the P3 and P4 residues are the primary determinants of the ability of a substrate to discriminate between t-PA and u-PA. Mutagenesis of the P3 and P4 residues of plasminogen activator inhibitor type 1, the primary physiological inhibitor of both t-PA and u-PA, confirmed this prediction and indicated a predominant role for the P3 residue. Appropriate replacement of both the P3 and P4 residues enhanced the t-PA specificity of plasminogen activator inhibitor type 1 by a factor of 600, and mutation of the P3 residue alone increased this selectivity by a factor of 170. These results demonstrate that the combination of substrate phage display and substrate subtraction methods can be used to discover specificity differences between very closely related enzymes and that this information can be utilized to create highly selective inhibitors.

The chymotrypsin family of serine proteases has evolved to include members with both widely divergent and intimately related substrate specificities (1). We chose two members of this family, tissue-type plasminogen activator (t-PA)[1] and urokinase (u-PA), to test the hypothesis that small molecule libraries could be used to identify substrates that discriminate between closely related enzymes. This choice of enzymes assured a rigorous test of the hypothesis because t-PA and u-PA possess an extremely high degree of structural similarity (2, 3), share the same primary physiological substrate (plasminogen) and inhibitors (plasminogen activator inhibitor types 1 and 2) (4, 5), and exhibit restricted substrate specificity (6–8).

Despite their striking similarities, the physiological roles of t-PA and u-PA are distinct (9), and many studies (10–16), including several that utilize transgenic mice (9, 11, 16), suggest that selective inhibition of either enzyme might have beneficial therapeutic effects. Mice lacking t-PA, for example, are resistant to specific excitotoxins that cause extensive neurodegeneration in wild type mice (11), and mice lacking u-PA exhibit defects in the proliferation and/or migration of smooth muscle cells in a model of restenosis following vascular injury (9). u-PA-deficient mice are also resistant to the induction and/or progression of several tumor types in a two-stage, chemical carcinogenesis model (16).

Because mice lacking either t-PA or u-PA do not develop thrombotic disorders, selective inhibition of either of these two enzymes seems unlikely to create thrombotic complications *in vivo*. On the other hand, mice lacking both t-PA and u-PA suffer severe thrombosis in many organs and tissues, resulting in a significantly reduced life expectancy (9). Nonselective inhibition of these two enzymes, therefore, seems almost certain to produce catastrophic consequences in the clinical setting. Consequently, significant interest exists in the development of inhibitors that are stringently specific for either t-PA or u-PA. Rational design of these selective inhibitors is greatly complicated, however, by the absence of obvious "lead compounds"; both their primary physiological substrate and inhibitors fail to discriminate between the two closely related proteases.

We have previously described the use of substrate phage display, a strategy originally developed by Matthews and Wells (17), to elucidate optimal subsite occupancy for substrates of t-PA and to isolate peptide substrates that were cleaved as much as 5300 times more efficiently by t-PA than peptides containing the primary sequence of the actual target site present in plasminogen (18). Four of these selected substrates, including the most labile t-PA substrate, were chosen for detailed characterization; consequently, small peptides containing the four selected amino acid sequences were synthesized, and a kinetic analysis of the cleavage of these peptides by t-PA was performed (18). Subsequent analysis of these four selected substrates demonstrated that they were also efficiently cleaved by u-PA and therefore were not selective for t-PA *versus* u-PA. Consequently, to facilitate the rapid isolation of peptide sub-

* This study was funded in part by National Institutes of Health Grants RO1 HL52475 and PO1 HL31950 (to E. L. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ An assistant investigator with the Howard Hughes Medical Institute. To whom correspondence may be addressed: Dept. of Pharmacology and Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, TX 75235.

** To whom correspondence may be addressed: Dept. of Vascular Biology (VB-1), The Scripps Research Institute, La Jolla, CA 92037. Fax: 619-784-7323.

[1] The abbreviations used are: t-PA, tissue-type plasminogen activator; u-PA, urokinase-type plasminogen activator; PAI-1, plasminogen activator inhibitor type 1; mAb, monoclonal antibody; fTC, fAFF1-tether C; HPLC, high pressure liquid chromatography.

strates that were cleaved at greater catalytic efficiencies by t-PA than by u-PA, we developed a novel protocol to prepare phage "substrate subtraction libraries" (Fig. 1). Characterization of individual members of the substrate subtraction library quickly accomplished the direct identification of highly labile, small peptide substrates that were preferentially cleaved by t-PA. In addition, insights gained during the analysis of these selective substrates were used to design a variant of the serpin PAI-1 whose selectivity toward t-PA was enhanced by a factor of 600.

MATERIALS AND METHODS

Reagents—Competent MC1061 (F⁻) *Escherichia coli* and nitrocellulose were purchased from Bio-Rad. Pansorbin (Protein A-bearing *Staphylococcus aureus*) cells were obtained from Calbiochem. K91 (F⁺) and MC1061 (F⁻) strains of *E. coli* were provided by Steve Cwirla (Affymax). mAb 179 which recognizes the epitope (ACLEPYTACD) of the human placental alkaline phosphatase protein with subnanomolar affinity, was provided by Ron Barrett (Affymax). mAb 3-E7 was purchased from Gramsch Laboratories (Schwabhausen, FRG). t-PA was obtained from Bruce Keyt (Genentech), and u-PA was obtained from Jack Henkin (Abbott Laboratories).

Construction of the phage vector fAFF1-tether C (fTC) and the random hexapeptide library fAFF-TC-LIB has been previously described (19). Control substrate phage fTC-P1, was constructed by hybridizing the single-stranded oligonucleotides 5′-TCGAGCGGTGGATCCGGTA-CTGGTCGTACTGGTCATGCCTCTGGTAC-3′ and 5′-CGCCACCTAGG-CCAGGACCAGCACAACAACCACGAGAC-3′ and then ligating the annealed, double-stranded products into the *XhoI*/*Kpn*I-cut vector fTC. All constructs were first transformed into MC1061 by electroporation and then transferred into K91.

Measurement of Enzyme Concentrations—Concentrations of functional t-PA and u-PA were measured by active site titration with 4-methylumbelliferyl *p*-guanidinobenzoate (20) using a Perkin-Elmer LS 50B luminescence fluorometer as described previously (7, 21). In addition, the enzyme were titrated with a standard PAI-1 preparation that had been previously titrated against a trypsin primary standard. Total enzyme concentrations were measured by enzyme-linked immunosorbent assay.

Phage Selection Using t-PA or u-PA—Conventional substrate phage display was originally developed by Matthews and Wells (17) using monovalent phage, and another method that used multivalent phage was reported later (19). Multivalent substrate phage were screened with t-PA using reaction conditions reported previously (18). Identical reaction conditions were used to screen the phage library with u-PA except that digestion of the phage was performed using enzyme concentrations varying from 5 to 10 μg/ml and incubation times varying from 1 to 10 h.

Preparation of Substrate Subtraction Libraries—The initial random hexapeptide library fAFF-TC-LIB was subjected to three rounds of high stringency screening with t-PA (18) to prepare an intermediate library containing phage whose randomized hexamer sequences were digested efficiently by t-PA. The intermediate library was then amplified and screened, at low stringency, with u-PA. Following digestion of the intermediate library with u-PA, mAb E-7 and immobilized protein A were added to the mixture, and the resulting ternary complexes were pelleted by centrifugation as described previously (18, 19). In contrast to all previous screening steps, however, we retained the precipitated ternary complexes and discarded the supernatant, which contained phage that were digested by u-PA. The precipitated ternary complexes were washed four times with 50 mM Tris (pH 7.5), 0.1 M NaCl, 1 mM EDTA, 0.01% Tween 80, resuspended, and digested with 50 μg/ml t-PA at 37 °C for 4 h. The remaining ternary complexes were precipitated by centrifugation and discarded, and the supernatant, which contained t-PA-selective phage, was retained. Phage in the final supernatant were amplified overnight in *E. coli* K91 cells. Following amplification, individual phage clones were functionally characterized using a dot blot assay and analyzed by DNA sequencing.

Dot Blot Assay of Phage Proteolysis—Phage precipitation and dot blot analysis were performed as described previously (18, 19). Individual phage stocks were prepared and digested with no enzyme, t-PA, u-PA, or u-PA in the presence of 1 mM amiloride, a specific inhibitor of u-PA, for periods of time varying from 15 min to 10 h. Individual reaction mixtures were spotted onto a nitrocellulose filter using a dot blotter apparatus (Bio-Rad). The filter was probed with mAb 3E-7 and developed using the Amersham Western ECL kit. Loss of positive staining indicates loss of antibody epitopes from the phage due to proteolytic cleavage of the randomized hexamer region.

Preparation and Sequencing of DNA from Phage Clones—DNA samples were prepared from interesting phage clones as described previously (18). Briefly, phage are precipitated from a 1-ml overnight culture by adding 200 μl of 20% polyethylene glycol in 2.5 M NaCl. The mixture was incubated on ice for 30 min, and the phage pellet was collected by microcentrifugation for 5 min. The phage were resuspended in 40 μl of lysis buffer (10 mM Tris-HCl, pH 7.6, 0.1 mM EDTA, 0.5% Triton X-100) and heated at 80 °C for 15 min. Single-stranded DNA was purified by phenol extraction and ethanol precipitation. One-third of the single-stranded DNA was used for dideoxy sequencing.

Kinetics of Cleavage of Synthetic Peptides by t-PA and u-PA—Peptides were synthesized and purified as described (7). Kinetic data were obtained by incubating various concentrations of peptide with a constant enzyme concentration to achieve between 5 and 20% cleavage of the peptide in each reaction. For assays with u-PA, enzyme concentration was either 815 or 635 nM. For assays with t-PA, enzyme concentration was 700 nM. Peptide concentrations were chosen where possible to surround K_m and in all cases were between 1 and 32 mM. The buffer used in these assays has been described (7). Reactions were stopped by the addition of trifluoroacetic acid to 0.33% or by freezing on dry ice. Cleavage of the 13- and 14-residue peptides was monitored by reverse phase HPLC as described (7). The 4–6-residue peptides were acylated at their amino termini and amidated at their carboxyl termini. Cleavage of the 4–6-residue peptides was monitored by hydrophilic interaction HPLC chromatography (22) using a polyhydroxyaspartamine column from PolyLC (Columbia, MD). Buffer A was 50 mM triethylamine phosphate in 10% acetonitrile, and buffer B was 10 mM triethylamine phosphate in 80% acetonitrile. Peptides were eluted by a gradient that was varied from 100% buffer B to 100% buffer A during a 13-min interval. The percentage of cleaved peptide was calculated by dividing the area under the product peaks by the total area under substrate and product peaks. For all peptides containing multiple basic residues, mass spectral analysis of products confirmed that cleavage occurred at a single site and identified the scissile bond. Data were interpreted by Eadie-Hofstee analysis. Errors were determined as described (23) and were <25%.

Site-directed Mutagenesis and Construction of Expression Vectors Encoding Variants of PAI-1—The expression vector pPAIS7/HS was derived from the plasmid pBR322 and contains a full-length cDNA encoding human PAI-1 that is transcribed from a T7 gene 10 promoter (24). The 300-base pair *SalI*/*Bam*HI fragment of human PAI-1 was subcloned from pPAIS7/HS into bacteriophage M13mp18. Single-stranded DNA produced by the recombinant M13mp18 constructs was used as a template for site-specific mutagenesis according to the method of Zoller and Smith (25) as modified by Kunkel (26).

Following mutagenesis, single-stranded DNA corresponding to the entire 300-base pair *SalI*-*Bam*HI fragment was fully sequenced to ensure the presence of the desired mutation and the absence of any additional mutation. The 300-base pair *SalI*-*Bam*HI double-stranded DNA fragments from mutated, replicative form DNAs were used to replace the corresponding fragment in pPAIS7/HS to yield full-length cDNAs encoding PAI/P3R, PAI/P4Q, and PAI/P4Q,P3R.

Expression and Purification of Recombinant PAI-1 Variants—Expression of wild type and mutated variants of PAI-1 was accomplished in the *E. coli* strain BL21[DE3]pLys^s (Novagen), which synthesizes T7 RNA polymerase in the presence of isopropyl-1-thio-β-D-galactopyranoside. Bacterial cultures were grown at 37 °C with vigorous shaking to A_{600} of 1.1–1.3, and isopropyl-1-thio-β-D-galactopyranoside was added to a final concentration of 1 mM to induce the synthesis of T7 RNA polymerase and the production of PAI-1 proteins. Cultures were grown for an additional 1–2 h at 37 °C and then shifted to 30 °C for 2–6 h.

Cells were pelleted by centrifugation at 8000 × *g* for 20 min at 4 °C and resuspended in 40 ml of cold start buffer (20 mM sodium acetate, 200 mM NaCl, and 0.01% Tween 20, pH 5.6). The cell suspension was disrupted in a French pressure cell (Aminco), and cellular debris was removed by ultracentrifugation for 25 min at 32,000 × *g*.

Purification of soluble, active PAI-1 was performed as described previously (27). PAI-1-containing supernatants were injected onto a XK-26 column (Pharmacia Biotech Inc.) packed with CM-50 Sephadex (Pharmacia). The column was washed with 5 column volumes of start buffer (20 mM sodium acetate, 200 mM NaCl, and 0.01% Tween 20, pH 5.6), and PAI-1 proteins were eluted using a 0.2–1.8 M linear gradient of NaCl in the same buffer. Peak fractions were collected, pooled, and concentrated using a centriplus 30 concentrator (Amicon). Purified preparations were analyzed by activity measurements using standard,

direct assays of t-PA, SDS-polyacrylamide gel electrophoresis, and measurement of optical density at 280 nm.

Measurement of Active PAI-1 in Purified Preparations—A primary standard of trypsin was prepared by active site titration using *p*-nitrophenyl *p'*-guanidinobenzoate HCl as described previously (28). Concentrations of active molecules in purified preparations of wild type or mutated PAI-1s were determined by titration of standardized trypsin as described by Olson *et al.* (29) and by titration of standardized t-PA preparations.

Kinetic Analysis of the Inhibition of t-PA and u-PA by Wild Type and Mutated Variants of PAI-1—Second order rate constants (k_i) for inhibition of tPA or uPA were determined using pseudo-first order ($k_i < 2 \times 10^6$) or second order ($k_i > 2 \times 10^6$) conditions. For each reaction, the concentrations of enzyme and inhibitor were chosen to yield several data points for which the residual enzymatic activity varied between 20 and 80% of the initial activity. Reaction conditions and data analysis for pseudo-first order reactions were as described previously (30–33).

For second order reactions, equimolar concentrations of u-PA and PAI-1 were mixed directly in microtiter plate wells and preincubated at room temperature for periods of time varying from 0 to 30 min. Following preincubation the mixtures were quenched with an excess of neutralizing anti-PAI-1 antibody (generously provided by Dr. David Loskutoff, The Scripps Research Institute), and residual enzymatic activity was measured using a standard, indirect chromogenic assay. These indirect, chromogenic assays were compared with control reactions containing no PAI-1 or with reactions to which PAI-1 was added after preincubation and the addition of anti-PAI-1 antibody, plasminogen, and Spec PL. Data were analyzed by plotting the reciprocal of the residual enzyme concentration *versus* the time of preincubation.

RESULTS AND DISCUSSION

Construction of Substrate Phage Libraries—A polyvalent fd phage library that displayed random hexapeptide sequences and contained 2×10^8 independent recombinants was prepared (18, 19). Each member of this library displayed an N-terminal extension from phage coat protein III containing a randomized region of six amino acids followed by a six-residue linker sequence (SSGGSG) and the epitopes for mAb 179 and mAb 3-E7. Because neither t-PA nor u-PA digests the protein III sequence, the antibody epitopes, or the flexible linker sequence, the loss of antibody epitopes from the phage surface upon incubation with either enzyme requires cleavage of the randomized peptide insert. Incubation of the library with t-PA, followed by removal of phage retaining the antibody epitopes, therefore, accomplishes the enrichment of phage clones whose random hexamer sequence can be cleaved by t-PA.

Construction of Substrate Subtraction Libraries—The initial phage library was subjected to three rounds of high stringency selection with t-PA to assure the preparation of an intermediate library that is highly enriched for phage that are efficient substrates of t-PA (Fig. 1). This intermediate library was then digested at low stringency with u-PA to remove phage that are moderate or good substrates for u-PA. Substrate subtraction was accomplished after the protease digestion of phage by adding mAb 3E-7 and immobilized protein A to the reaction mixture and precipitating the ternary complexes that contain the undigested phage. By contrast to all earlier selections, the phage remaining in solution were discarded, and the precipitate containing the ternary complexes was resuspended. Phage that were preferentially cleaved by t-PA were then released from the ternary complexes by digestion with t-PA.

Identification and Kinetic Characterization of t-PA-selective Substrates—Using the protocol outlined in Fig. 1 and a previously described (18, 19) sensitive dot blot assay (Fig. 2), we isolated and functionally verified 37 t-PA-selective phage clones that contained 32 distinct substrate sequences (Fig. 3). As illustrated by the experiment depicted in Fig. 2, the dot blot assay can rapidly provide information regarding both the activity and specificity of individual substrate phage clones. Based on the results of these assays, three peptide substrates (II–IV) containing hexamer sequences present in individual

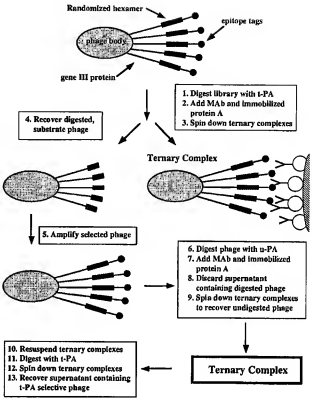

FIG. 1. Outline of the protocol used to create substrate subtraction libraries. The gene III fusion protein, phage, monoclonal antibodies, and immobilized protein A are not drawn to scale.

	PL	7	35	51
No Enzyme	●	●	●	●
tPA	●			●
uPA + Amiloride	●	●	●	●
uPA	●	●		

FIG. 2. Functional analysis of individual control or substrate phage stocks using a dot blot assay that has been previously described (18, 19). Loss of positive staining indicated removal of antibody epitopes from the phage by proteolytic cleavage of the random hexamer region. Control phage PL contained the P3–P3' region of the actual target sequence present in plasminogen (PGRVVG) and was not digested by either enzyme under the conditions used in this experiment. Substrate phage 51 was isolated from the intermediate library used to create the t-PA substrate subtraction library, contained the hexamer RIARRA, and was an efficient substrate of both t-PA and u-PA. Phage 7 was a member of the t-PA subtraction library, contained the hexamer FRGRAA, and, as predicted, was a t-PA-selective substrate. Phage 35 was isolated from a conventional u-PA substrate library, contained the hexamer RSANAI, and was a u-PA-selective substrate.

members of the substrate subtraction library were synthesized and characterized to provide a quantitative analysis of the properties of putative t-PA-selective substrates. These peptides were cleaved 180–1500-fold times more efficiently by t-PA than a control peptide (I) containing the physiological cleavage site present in plasminogen (Table I). In addition, by contrast to the plasminogen-derived, control peptide, which was cleaved 3-fold more efficiently by u-PA than by t-PA, the three selected peptides were cleaved 13–47 times more efficiently by t-PA than by u-PA.

The t-PA/u-PA selectivity possessed by members of the substrate subtraction library was significantly greater than that displayed by members of an "optimized" substrate library constructed using standard substrate phage display methods. We

Clone	P5	P4	P3	P2	P1	P'1	P'2	P'3
1		A	L	R	R	G	D	
2	D	Y	R	G	R	M	(L)	
3		E	R	A	R	G	A	
4		E	R	L	R	K	A	
5		(G)	F	G	R	H	A	A
6		F	L	P	R	T	A	
7		F	R	G	R	A	A	
8		H	R	M	R	M	G	
9		H	Y	G	R	S	G	
10		(G)	I	M	R	R	G	K
11	I	T	Y	G	R	R	(L)	
12		K	F	T	R	S	G	A
13		L	I	P	R	R	A	A
14	M	T	R	K	R	M	(L)	
15		N	F	A	R	M	G	
16		N	H	L	R	K	A	
17		N	V	G	R	M	G	
18		N	V	S	R	R	G	
19		P	I	S	R	R	A	
20		P	V	G	R	M	G	
21		Q	R	G	R	K	A	
22		R	L	L	R	S	V	
23		S	F	G	R	R	H	
24	S	L	R	G	R	S	(L)	
25		T	V	L	R	R	A	
26		(G)	V	A	R	R	V	K
27		V	I	A	R	S	N	
28		V	N	T	K	S	G	
29		V	R	A	R	G	A	
30	V	R	R	G	R	S	(L)	
31		V	R	R	R	G	A	K
32		T	R	V	R	A	K	

FIG. 3. **Primary sequence of the randomized hexamer found in phage that were selective substrates of t-PA.** Peptide sequences have been shifted to the left or right to align corresponding subsites of each individual hexamer sequence. Amino acids in *parenthesis* are flanking residues from the gene III fusion protein.

have previously presented detailed, kinetic analysis of the cleavage by t-PA of four substrates from such an optimized library (18). Subsequent analysis, using the dot blot assay, revealed that these four selected substrates were also cleaved efficiently by u-PA and therefore were not highly selective for t-PA *versus* u-PA. In fact, one of these substrates, SRARKA, was actually cleaved more rapidly by u-PA than by t-PA. The remaining three selected substrates were cleaved slightly more rapidly in the dot blot assay by t-PA than by u-PA. To determine the precise extent of selectivity toward t-PA exhibited by these substrates, we measured the catalytic efficiency of both t-PA and u-PA for hydrolysis of peptides containing these three selected amino acid sequences (Table I, peptides V–VII). In contrast to the 13–47-fold t-PA/u-PA selectivity exhibited by substrates isolated from the subtraction library, the t-PA/u-PA selectivity of these substrates identified using standard substrate phage display was 2.5–4.9-fold. These data verify the value of the subtraction library and indicate that, in this instance, a single subtraction step resulted in the recovery of substrates whose selectivity was enhanced by approximately 10-fold.

Comparison of Consensus Sequences Derived Using Substrate Phage Display and Substrate Subtraction Methods—We examined the consensus sequences derived using both standard and subtractive phage display to identify determinants of substrates that mediated t-PA/u-PA selectivity. Using standard substrate phage display protocols, we inferred the consensus sequence GR↓*X*A, where *X* could be a number of different amino acids but most often was arginine, for optimal substrates of t-PA (18). Strong conservation of particular amino acids at the P3 position of substrates was not observed in these experiments; however, slightly more than half (57%) of the selected substrates did contain one of five hydrophobic residues (Leu, Ile, Val, Phe, Tyr) at this position. Similarly, no obvious conservation of any particular amino acid was observed at the P4 position of the optimized substrates, although the majority (60%) contained one of four small residues (Ser, Thr, Gly, Ala) at this position (18).

To assure that members of the substrate subtraction library were highly active, as well as unusually specific, toward t-PA, the substrate subtraction library was first subjected to high stringency selection using t-PA. Consequently, it is not surprising that substrates obtained from the subtraction library agree well with the GR↓*X*A consensus sequence observed in the optimized library. Significant differences were observed, however, when comparing residues found at the P3 and P4 positions of substrates from the two libraries. 100% of the substrates from the subtraction library contained a large residue at P3, and 97% of the sequences contained either arginine or a large hydrophobic residue at this position. Arginine was the most frequently observed individual residue at P3 in the subtraction library, occurring in approximately 38% of the selected sequences. In addition, while 60% of the sequences in the optimized library contained small residues at P4, 75% of the sequences from the subtraction library contained a large residue at this position. Comparison of these consensus sequences, therefore, suggested that the P3 and P4 residues were key determinants of enhanced t-PA/u-PA selectivity, with preferences for arginine at P3 and large hydrophobic residues at P4.

Important Distinctions between Substrate Phage Display and Substrate Subtraction Techniques—A key distinction between substrate subtraction libraries and substrate phage libraries obtained using conventional protocols is that the former are actively influenced by both positive and negative determinants of specificity. While substrate phage display identifies the most labile substrates for a particular enzyme, substrate subtraction identifies the most selective substrates for the enzyme.

With enzymes for which the most labile substrates are also selective substrates (*e.g.* u-PA),[2] the two protocols may yield similar results. Even in this situation, however, both techniques remain useful because it seems unlikely that the standard protocols will yield the most selective substrate or that substrate subtraction will yield the most labile substrate. In contrast, with enzymes like t-PA, where the most labile substrates are not also highly selective substrates, the two methods produce distinct results. For example, the most labile substrate sequence identified for t-PA, PFGRSA, was isolated by standard substrate display protocols and not by substrate subtraction. This sequence was also cleaved efficiently by u-PA (k_{cat}/K_m = 320 M^{-1} s^{-1}) and therefore exhibited only 4.7-fold t-PA/u-PA selectivity. This observation emphasizes the importance of subtle determinants of substrate specificity and the ability of subtraction libraries to identify combinations of amino acids that maximize selectivity. The most selective P4-P1 sequence, FRGR (Table I, peptide X), occurred in the substrate subtraction library but not in the substrate phage library, although the later protocol included one additional round of screening and approximately 40% more positive phage.

For reasons described above, substrate subtraction techniques greatly facilitate elucidation of subsite occupancy that contributes primarily to the specificity, rather than to the catalytic efficiency, of substrate hydrolyses. Data presented in this study suggest that appropriate occupancy of P3 for t-PA substrates is one example of this type of interaction. In physiological contexts, such subtle specificity determinants may be essential. For example, *Erythrina* trypsin inhibitor inhibits t-PA but not u-PA. Although the molecular basis of this specificity

[2] S. H. Ke, G. S. Coombs, K. Tachias, D. R. Corey, and E. L. Madison, unpublished observations.

TABLE I
Comparison of k_{cat}, K_m, and k_{cat}/K_m for the hydrolysis by t-PA or u-PA of peptides selected using substrate subtraction or standard substrate phage display protocols

	Substrate (P$_n$. . ., P2, P1 ↓ P1′, P2′, . . . P$_n$′)*	t-PA			u-PA			t-PA/u-PA selectivity
		k_{cat}	K_m	k_{cat}/K_m	k_{cat}	K_m	k_{cat}/K_m	
		s^{-1}	μM	$M^{-1}\ s^{-1}$	s^{-1}	μM	$M^{-1}\ s^{-1}$	
	Native cleavage sequence from plasminogen							
I	KKSPGR ↓ VVGGSVAH	0.0043	15,000	0.29	0.003	3400	0.88	0.33
	Peptides from substrate subtraction library							
II	LGGSGQRGR ↓ KALE	0.99	2300	430	0.02	2180	9.2	47
III	LGGSGERAR ↓ GALE	0.073	1410	52	0.004	970	4.0	13
IV	LGGSGHYGR ↓ SGLE	1.29	4010	322	0.059	3800	15	21
	Peptides from standard substrate phage library							
V	GGSGWLGR ↓ RGLVPE	2.0	10,000	200	0.32	4000	80	2.5
VI	GGSGYTGR ↓ RGLVPE	1.6	7300	220	0.20	4400	45	4.9
VII	GSGPPGR ↓ SALVPE	3.3	2200	1500	0.71	2200	320	4.7
	Minimized peptides from substrate subtraction library							
VIII	YGR ↓ S	23.7	6000	3950	2.6	11,400	230	17
IX	RGR ↓ K	15.3	16,600	922	0.76	46,500	16	57
X	FRGR ↓ K	12.2	9800	1240	0.14	8600	16	78

* Positional nomenclature of subsite residues. Arrows denote the position of peptide bond hydrolysis. The peptide bond is cleaved between P1 and P1′. The S.E. in these determinations was 4–22%.

remains obscure, it is intriguing that the *Erythrina* inhibitor contains an arginine in the P3 position (34).

Although standard substrate phage display protocols can be used successfully to isolate selective substrates, this task will frequently prove arduous due to the necessity of isolating, sequencing, and functionally characterizing an inordinately large number of phage. We find, however, that a straightforward modification of these protocols to include a subtraction step substantially enhances the efficiency of this process and therefore significantly increases the rate at which highly selective substrates can be discovered. For example, we performed two independent substrate phage display experiments with t-PA using standard protocols. Characterization by dot blot analysis of a total of 105 individual phage clones isolated in the final round of selection in the two experiments revealed that either 6.7% (first experiment) or 4.2% (second experiment) of these phage clones were t-PA-selective substrates. In contrast, characterization of 71 phage clones from a substrate subtraction library indicated that 54.9% of these phage were t-PA-selective substrates. In other words, depending on the protocol being used, we were forced to isolate, prepare, and characterize, on average, either 15–24 clones (standard substrate display) or 1.8 clones (substrate subtraction) to obtain a single t-PA-selective phage.

Minimization of the Selective Peptide Substrates—The kinetic analysis described above was performed using substrate peptides that were 14 amino acids in length. To confirm that the specificity we observed was inherent in the selected hexapeptide sequences, we examined the kinetics of cleavage of short peptides containing only sequences found within selected hexapeptide sequences. Tetrapeptide IX and pentapeptide X, for example, were cleaved 57 or 78 times, respectively, more efficiently by t-PA than by u-PA and were therefore actually more selective than the 14-mer peptides (II–IV). Furthermore, tetrapeptide IX and pentapeptide X were significantly more labile substrates of t-PA than any of the 14-mer peptides (II–IV). Thus, compared with the 14-mers, the tetra- and pentapeptides not only maintained specificity but also acquired increased activity. These data confirm the proposed status of the P3 and P4 residues as specificity determinants for substrates of t-PA and u-PA, suggest a particularly prominent role of the P3 residue in this capacity, and demonstrate that highly selective substrates of t-PA can be created by appropriate occupancy of the P3–P1′ subsites alone. These observations may also provide a firm basis for the rational design of highly selective, small molecule inhibitors of t-PA. Although hydrolysis of the selective, small peptide substrates by t-PA is characterized by K_m values in the millimolar range, it has been routinely observed that the introduction of a transition state bond geometry adjacent to the P1 residue of a protease substrate can create either a reversible inhibitor whose affinity for the target protease is enhanced by 3–6 orders of magnitude or an irreversible inhibitor with an impressive second order rate constant for inhibition of the target protease ($>10^5\ M^{-1}\ s^{-1}$) (for a review, see Ref. 35). Similar results using the substrates identified in this study would create highly selective t-PA inhibitors, with affinities in the low nanomolar range, that might be further improved by subsequent, systematic chemical modification.

Design and Characterization of Variants of PAI-1 That Are Selective for t-PA—To test the prediction, based on analysis of the cleavage of peptide substrates, that the P3 residue can mediate the ability of an inhibitor to discriminate between t-PA and u-PA, we performed site-specific mutagenesis of PAI-1, the primary physiological inhibitor of both t-PA and u-PA. Three variants of PAI-1 were produced and characterized: a variant in which the P3 serine[3] was converted to an arginine residue, a variant in which the P4 valine was replaced by a glutamine residue, and a double mutant containing both of these substitutions. Kinetic analysis of the inhibition of t-PA and u-PA by these variants of PAI-1 proved consistent with conclusions drawn from the previous experiments utilizing peptide substrates. The second order rate constants for inhibition of t-PA and u-PA by wild type PAI-1 were $1.6 \times 10^6\ M^{-1}\ s^{-1}$ and $1.9 \times 10^7\ M^{-1}\ s^{-1}$, respectively (Table II). Thus, wild type PAI-1 exhibits approximately 11.9-fold specificity toward u-PA. By contrast, the second order rate constants for inhibition of t-PA and u-PA by the P3 arginine variant of PAI-1 were, respectively, $1.4 \times 10^6\ M^{-1}\ s^{-1}$ and $1.0 \times 10^5\ M^{-1}\ s^{-1}$, an approximately 170-fold reversal in specificity (Table II). This large alteration in specificity was achieved without sacrificing activity toward the target enzyme; the P3 arginine mutation reduced activity of PAI-1 toward u-PA by a factor of approximately 190 without significantly affecting reactivity toward t-PA.

Individual mutation of the P4 valine of wild type PAI-1 to a glutamine residue had no effect on the rate of inhibition of either t-PA or u-PA. As suggested by the predominance in the subtraction library of substrates containing both large P3 and large P4 residues, however, the P4 glutamine mutation did

[3] According to the PAI-1 numbering system, the P3 residue is serine 344, and the P4 residue is valine 343.

TABLE II
Second order rate constants for inhibition of t-PA or u-PA by wild type and variants of PAI-1

Inhibitor	Primary sequence of reactive center loop (P4–P2′)	Rate constant toward t-PA	Rate constant toward u-PA	t-PA/u-PA selectivity
		$M^{-1}\ s^{-1}$	$M^{-1}\ s^{-1}$	
Wild type PAI-1	VSAR↓MA	1.6×10^6	1.9×10^7	0.08
PAI-1/P3R	VRAR↓MA	1.4×10^6	1.0×10^5	14
PAI-1/P4Q	QSAR↓MA	1.6×10^6	1.9×10^7	0.08
PAI-1/P4Q,P3R	QRAR↓MA	1.4×10^6	2.9×10^4	48

increase the t-PA selectivity of the P3 arginine variant of PAI-1. The second order rate constants for the inhibition of t-PA and u-PA by the P4 glutamine, P3 arginine double mutant of PAI-1 were $1.4 \times 10^6\ M^{-1}\ s^{-1}$ and $2.9 \times 10^4\ M^{-1}\ s^{-1}$, respectively (Table II). While maintaining full activity toward t-PA, then, the double mutant exhibited an approximately 600-fold enhanced t-PA/u-PA selectivity compared with wild type PAI-1 and a 3.5-fold greater t-PA selectivity than the P3 arginine variant of PAI-1. The absolute t-PA/u-PA selectivity of wild type PAI-1, the P3 Arg single mutant, and the P3 Arg, P4 Gln double mutant was approximately 0.08, 14, and 48, respectively.

Contribution of Structural Studies to Understanding Restricted Specificity at P2 and the Critical Role of the P3 and P4 Residues in Mediating Specificity—At the time these studies were initiated, no structural information was available for the protease domain of either t-PA or u-PA. During preparation of this manuscript, however, both of these important structures were described (2, 3). The t-PA structure contained a benzamidine molecule, which occupied the S1 pocket, and the u-PA structure included a covalently bound, tripeptide chloromethyl ketone that filled the S1, S2, and S3 sites. These structural reports have provided a clear molecular basis for our observation that both t-PA and u-PA display a strong preference for glycine at the P2 subsite of substrates. In both enzymes, the side chain of residue 99 (chymotrypsin numbering), a tyrosine in t-PA and a histidine in u-PA, severely restricts the size of the S2 pocket. Modeling of even an alanine residue into this subsite produces a slight steric conflict with residue 99 of the enzymes, and larger residues at P2 create extensive clashes.

Another important observation of the structural studies is that the major distinction between the substrate binding cleft of t-PA and u-PA occurs in the region corresponding to the aryl binding site of thrombin (2, 3, 36). In u-PA this pocket is partially filled by an insertion of two amino acids (threonine 97A, leucine 97B; chymotrypsin numbering) that is absent in t-PA. Consequently, the aryl biding site is significantly larger in t-PA than in u-PA. In addition, unlike Ser^{174} of u-PA, Arg^{174} of t-PA extends toward, and partially occupies, the aryl binding site. However, because Arg^{174} appears highly flexible and mobile (2), the extent to which this residue actually influences the aryl binding site of t-PA in solution remains uncertain.

Depending on the precise binding mode, the aryl binding site can interact with any of several residues of a particular substrate or inhibitor. If the substrate or inhibitor adopts a canonical conformation within the active site cleft, the P4 residue will occupy the aryl binding site (37). However, the most extensively studied aryl binding site, that of thrombin, interacts with the P9 residue of the physiological substrate fibrinogen (38). In addition, the P2 glycine of substrates selected in this study, with its greatly expanded range of allowed conformations compared with other natural amino acids, could allow the P3 residue to interact with the aryl binding site. Consequently, the structural studies indicating that the major differences between the active site cleft of the two proteases occur in the aryl binding site are consistent with our substrate phage studies, which demonstrate that the P3 and P4 residues are the primary determinants of the ability of a substrate to distinguish t-PA and u-PA.

Differences at position 217, a leucine in t-PA and an arginine in u-PA, may also contribute to specificity differences observed in this study for the two enzymes. In the u-PA structure, Arg^{217} adopts an unusual conformation that allows formation of a salt bridge with the P3 glutamic acid of the bound inhibitor. Clearly, a similar juxtaposition of Arg^{217} and the P3 arginine residue of a substrate or inhibitor would create an unfavorable electrostatic interaction that would be absent when the same substrate or inhibitor interacted with t-PA. However, if Arg^{217} of u-PA adopted a more commonly observed conformation, this residue would extend into solvent and could be located at a significant distance from the P3 arginine residue. Thus, the extent to which Arg^{217} of u-PA and Leu^{217} of t-PA contribute to specificity distinctions between the two enzymes remains an open question.

As discussed above, insights gained from this study and from very recent structural studies do allow the identification of candidate residues that may mediate important specificity distinctions between t-PA and u-PA. The absence of key structural information regarding the binding mode(s) of a particular selective substrate or inhibitor to both t-PA and u-PA, however, precludes a definitive description at atomic resolution of the mechanisms by which the specificity observed in our studies has been achieved.

Limitations of Structural Studies and Contribution of Substrate Phage and Substrate Subtraction Techniques to Understanding Protease Specificity—It will not be possible to understand enzyme catalysis and specificity without extensive structural information. On the other hand, it is very unlikely that the recent structural studies alone could have been used to predict the most selective substrates of either t-PA or u-PA. One difficulty described above is that the structural data for the proteases does not necessarily provide information regarding the precise binding mode and backbone conformation of an individual substrate, a critical detail that determines which, if any, region of the substrate will interact with specific regions (*e.g.* the aryl binding site) of the enzyme.

Another limitation is that the static structures do not necessarily provide information regarding the role of molecular dynamics in enzyme catalysis and specificity. Consequently, contributions to substrate specificity arising from subtle differences in dynamic properties of related enzymes may not be detected by the x-ray structures. The development of a detailed, molecular understanding of enzyme mechanism will therefore require information from studies that utilize a wide variety of techniques of molecular biology, biochemistry, and biophysics. Substrate phage display and substrate subtraction libraries can contribute to these efforts by providing a powerful, combinatorial approach to the identification of key determinants of substrate reactivity and specificity.

Conclusion—The rational design of small molecule inhibitors as therapeutic agents is often complicated by the necessity of discriminating between closely related enzymes (39). We dem-

onstrate here that appropriate selections of substrate phage can achieve this discrimination. Substrate subtraction libraries are likely to provide substrates that can distinguish between any two distinct proteases, and there is no theoretical reason why multiple proteases could not be used in the subtraction step to achieve even greater specificity. Moreover, it should be possible to prepare both substrate and substrate subtraction libraries as described above for any enzymes that can use peptides or proteins as substrates. These techniques might be adapted to protein kinases, for example, by using antibodies against phosphoserine, phosphothreonine, or phosphotyrosine during the selection of substrate phage. Consequently, the construction and characterization of substrate and substrate subtraction libraries may make substantial contributions to the rational design of highly specific, small molecule inhibitors of selected enzymes, a problem of paramount importance during the development of new therapeutic agents. In addition, by revealing specificity determinants that might otherwise remain obscure, these libraries will provide key insights into the molecular basis of specificity for a variety of important enzymes.

Acknowledgments—We thank Drs. Guy Salvesen, Steve Kent, Dave Loskutoff, and Mark Ginsberg for stimulating discussions and critical review of this manuscript. We thank Lynn Mayfield, Bikash C. Pramanik, and Dr. Clive Slaughter for peptide synthesis and mass spectral analysis and Karen Barker for initial kinetic analysis of the peptide substrates.

REFERENCES

1. Perona, J. J., and Craik, C. S. (1995) *Protein Sci.* 4, 337–360
2. Lamba, B., Bauer, M., Huber, R., Fischer, S., Rudolph, R., Kohnert, U., and Bode, W. (1996) *J. Mol. Biol.* 258, 117–135
3. Spraggon, G., Phillips, C., Nowak, U. K., Ponting, C. P., Saunders, D., Dobson, C. M., Stuart, D. I., and Jones, E. Y. (1995) *Structure* 3, 681–691
4. Callen, D., and Lijnen, H. R. (1991) *Blood* 78, 3114–3124
5. Madison, E. L. (1994) *Fibrinolysis* 8, Suppl. 1, 221–236
6. Ganu, V. S., and Shaw, E. (1982) *Peptide Protein Res.* 20, 421–428
7. Madison, E. L., Coombs, G. S., and Corey, D. R. (1995) *J. Biol. Chem.* 270, 7558–7562
8. Coombs, G. S., Dang, A. T., Madison, E. L., and Corey, D. R. (1996) *J. Biol. Chem.* 271, 4461–4467
9. Carmeliet, P., Schoonjans, L., Kieckens, L., Ream, B., Degen, J., Bronson, R., De Vos, R., van den Oord, J. J., Collen, D., and Mulligan, R. C. (1994) *Nature* 368, 419–424
10. Dano, K., Andreasen, P. A., Grondahl-Hansen, J., Kristensen, P., Nielsen, L. S., and Skriver, L. (1985) *Adv. Cancer Res.* 44, 139–266
11. Tsirka, S. E., Gualandris, A., Amaral, D. G., and Strickland, S. (1995) *Nature* 377, 340–344
12. Fukudome, K., and Esmon, C. T. (1994) *J. Biol. Chem.* 269, 26486–26491
13. Min, H. Y., Doyle, L. V., Vitt, C. R., Zandonella, C. L., Stratton-Thomas, J. R., Shuman, M. A., and Rosenberg, S. (1996) *Cancer Res.* 56, 2428–2433
14. Seeds, N. W., Williams, B. L., and Bickford, P. C. (1995) *Science* 270, 1992–1994
15. Osowski, L. (1988) *Cell* 52, 321–328
16. Shapiro, R. L., Duquette, J. G., Roses, D. F., Nunes, I., Harris, M. N., Kamino, H., Wilson, E. L., and Rifkin, D. B. (1996) *Cancer Res.* 56, 3597–3604
17. Matthews, D. J., and Wells, J. A. (1993) *Science* 260, 1113–1117
18. Ding, L., Coombs, G. S., Strandberg, L., Navre, M., Corey, D. R., and Madison, E. L. (1995) *Proc. Natl. Acad. Sci. U. S. A.* 92, 7627–7631
19. Smith, M. M., Shi, L., and Navre, M. (1995) *J. Biol. Chem.* 270, 6440–6449
20. Jameson, G. W., Roberts, D. V., Adams, R. W., Kyle, S. A., and Elmore, D. T. (1973) *Biochem. J.* 131, 107–117
21. Urano, T., Urano, S., and Castellino, F. J. (1988) *Biochem. Biophys. Res. Commun.* 150, 45–51
22. Alpert, A. J. (1990) *J. Chromatogr.* 499, 177–196
23. Taylor, J. R. (1982) *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*, University Science Books, Mill Valley, CA
24. Tucker, H. M., Mottonen, J., Goldsmith, E. J., and Gerard, R. D. (1995) *Nat. Struct. Biol.* 2, 442–445
25. Zoller, M. J., and Smith, M. (1984) *DNA* 3, 479–488
26. Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U. S. A.* 82, 488–492
27. Sancho, E., Tonge, D. W., Hockney, R. C., and Booth, N. A. (1994) *Eur. J. Biochem.* 224, 125–134
28. Chase, T., and Shaw, E. (1967) *Biochem. Biophys. Res. Commun.* 29, 508–514
29. Olson, S. T., Bock, P. E., Kvassman, J., Shore, J. D., Lawrence, D. A., Ginsburg, D., and Björk, I. (1995) *J. Biol. Chem.* 270, 30007–30017
30. Holmes, W. E., Lijnen, H. R., and Collen, D. (1987) *Biochemistry* 26, 5133–5140
31. Beatty, K., Bieth, J., and Travis, J. (1980) *J. Biol. Chem.* 255, 3931–3934
32. Madison, E. L., Goldsmith, E. J., Gerard, R. D., Gething, M. J., Sambrook, J. F., and Bassel-Duby, R. S. (1990) *Proc. Natl. Acad. Sci. U. S. A.* 87, 3530–3533
33. Madison, E. L., and Sambrook, J. F. (1993) *Methods Enzymol.* 223, 249–271
34. Teixeira, A. V., Dowdle, E. B. D., and Botes, D. P. (1994) *Biochim. Biophys. Acta* 1217, 23–28
35. Baggio, R., Shi, Y.-Q., Wu, Y.-Q., and Abeles, R. H. (1996) *Biochemistry* 35, 3551–3553
36. Bode, W., Turk, D., and Sturzebecher, J. (1990) *Eur. J. Biochem.* 193, 175–183
37. Bode, W., Turk, D., and Karshikov, A. (1992) *Protein Sci.* 1, 426–471
38. Ni, F., Meinwald, Y. C., Vásquez, M., and Scheraga, H. A. (1989) *Biochemistry* 28, 3094–3105
39. Crooke, S. T. (1996) *Nat. Biotechnol.* 14, 238–241